Connect. Accelerate. Outperform.™

Mellanox OFED for Linux Release Notes

Rev 2.2-1.0.1

Last Modified: July 13, 2014

www.mellanox.com

NOTE:
THIS HARDWARE, SOFTWARE OR TEST SUITE PRODUCT (“PRODUCT(S)”) AND ITS RELATED DOCUMENTATION ARE PROVIDED BY MELLANOX TECHNOLOGIES “AS-IS” WITH ALL FAULTS OF ANY KIND AND SOLELY FOR THE PURPOSE OF AIDING THE CUSTOMER IN TESTING APPLICATIONS THAT USE THE PRODUCTS IN DESIGNATED SOLUTIONS. THE CUSTOMER'S MANUFACTURING TEST ENVIRONMENT HAS NOT MET THE STANDARDS SET BY MELLANOX TECHNOLOGIES TO FULLY QUALIFY THE PRODUCTO(S) AND/OR THE SYSTEM USING IT. THEREFORE, MELLANOX TECHNOLOGIES CANNOT AND DOES NOT GUARANTEE OR WARRANT THAT THE PRODUCTS WILL OPERATE WITH THE HIGHEST QUALITY. ANY EXPRESS OR IMPLIED WARRANTIES, INCLUDING, BUT NOT LIMITED TO, THE IMPLIED WARRANTIES OF MERCHANTABILITY, FITNESS FOR A PARTICULAR PURPOSE AND NONINFRINGEMENT ARE DISCLAIMED. IN NO EVENT SHALL MELLANOX BE LIABLE TO CUSTOMER OR ANY THIRD PARTIES FOR ANY DIRECT, INDIRECT, SPECIAL, EXEMPLARY, OR CONSEQUENTIAL DAMAGES OF ANY KIND (INCLUDING, BUT NOT LIMITED TO, PAYMENT FOR PROCUREMENT OF SUBSTITUTE GOODS OR SERVICES; LOSS OF USE, DATA, OR PROFITS; OR BUSINESS INTERRUPTION) HOWEVER CAUSED AND ON ANY THEORY OF LIABILITY, WHETHER IN CONTRACT, STRICT LIABILITY, OR TORT (INCLUDING NEGLIGENCE OR OTHERWISE) ARISING IN ANY WAY FROM THE USE OF THE PRODUCT(S) AND RELATED DOCUMENTATION EVEN IF ADVISED OF THE POSSIBILITY OF SUCH DAMAGE.

Mellanox Technologies
350 Oakmead Parkway Suite 100
Sunnyvale, CA 94085
U.S.A.
www.mellanox.com
Tel: (408) 970-3400
Fax: (408) 970-3403

Mellanox Technologies, Ltd.
Beit Mellanox
PO Box 586 Yokneam 20692
Israel
www.mellanox.com
Tel: +972 (0)74 723 7200
Fax: +972 (0)4 959 3245

© Copyright 2014. Mellanox Technologies. All Rights Reserved.

Mellanox®, Mellanox logo, BridgeX®, ConnectX®, Connect-IB®, CoolBox®, CORE-Direct®, InfiniBridge®, InfiniHost®, InfiniScale®, MetroX®, MLNX-OS®, PhyX®, ScalableHPC®, SwitchX®, UFM®, Virtual Protocol Interconnect® and Voltaire® are registered trademarks of Mellanox Technologies, Ltd.

ExtendX™, FabricIT™, Mellanox Open Ethernet™, Mellanox Virtual Modular Switch™, MetroDX™, TestX™, Unbreakable-Link™ are trademarks of Mellanox Technologies, Ltd.

All other trademarks are property of their respective owners.

Table of Contents

Table of Contents 3
List Of Tables 4
Chapter 1 Overview 5
Chapter 2 Main Features in This Release 5
Chapter 3 Content of Mellanox OFED for Linux 6
Chapter 4 Supported Platforms and Operating Systems 7
4.1 Supported Hypervisors 7
Chapter 5 Hardware and Software Requirements 8
Chapter 6 Supported HCAs 8
Chapter 7 Compatibility 9
Chapter 8 Change Log History 10
8.1 Changes in Rev 2.2-1.0.1 From Rev 2.1-1.0.6 10
8.2 Changes in Rev 2.1-1.0.6 From Rev 2.1-1.0.0 10
8.3 Changes in Rev 2.1-1.0.0 From Rev 2.0-3.0.0 11
8.4 Changes in Rev 2.0-3.0.0 From Rev 2.0-2.0.5 12
8.5 New Features in Rev 2.0-2.0.5 12
Chapter 9 Known Issues 14
Chapter 10 API Changes 26
10.1 API Changes in MLNX_OFED Rev 2.2-1.0.1 26
10.2 API Changes in MLNX_OFED Rev 2.1-1.0.0 27
10.3 API Changes in MLNX_OFED Rev 2.0-3.0.0 27
10.4 API Changes in MLNX_OFED Rev 2.0-2.0.5 28
Chapter 11 Bug Fixes History 29

List Of Tables

Table 1: Mellanox OFED for Linux Software Components 6
Table 2: Supported Platforms and Operating Systems 7
Table 3: Additional Software Packages. 8
Table 4: MLNX_OFED Rev 2.2-1.0.1 Compatibility Matrix 9
Table 5: New Features, Changes and Fixes in v2.2-1.0.1 10
Table 6: New Features, Changes and Fixes in v2.1-1.0.6 10
Table 7: New Features, Changes and Fixes in v2.1-1.0.0 11
Table 8: New Features, Changes and Fixes in v2.0-3.0.0 12
Table 9: Known Issues. 14
Table 10: API Changes in MLNX_OFED Rev 2.2-1.0.1 26
Table 11: API Changes in MLNX_OFED Rev 2.1-1.0.0 27
Table 12: API Changes in MLNX_OFED Rev 2.0-3.0.0 27
Table 13: API Changes in MLNX_OFED Rev 2.0-2.0.5 28
Table 14: Fixed Bugs List 29

1 Overview

These are the release notes of Mellanox OFED for Linux Driver, Rev 2.2-1.0.1. Mellanox OFED is a single Virtual Protocol Interconnect (VPI) software stack and operates across all Mellanox network adapter solutions supporting the following uplinks to servers:

- 10, 20, 40 and 56 Gb/s InfiniBand (IB)
- 10, 40 and 56[1] Gb/s Ethernet
- 2.5 or 5.0 GT/s PCI Express 2.0
- 8 GT/s PCI Express 3.0

2 Main Features in This Release

MLNX_OFED Rev 2.2-1.0.1 provides the following new features:

- Signature Verbs (T10-PI)
- Connect-IB® DC support
- Single port Virtual Function
- Multi-GUID support for InfiniBand in SR-IOV
- MSI-X vector allocation for the SR-IOV hypervisor and guest
- Ethernet VXLAN Hardware Stateless offload (ConnectX®-3 Pro ONLY)
- Ethernet Power Management Quality of Service
- Ethernet PTP Hardware Clock support
- QP0 query operations for privileged Virtual Functions
- OVS support for eIPoIB
- Support for up to 126 Virtual Functions (pending OS and Server support)

1. 56 GbE is a Mellanox propriety link speed and can be achieved while connected to Mellanox SX10XX switch series

3 Content of Mellanox OFED for Linux

Mellanox OFED for Linux software contains the following components:

Table 1 - Mellanox OFED for Linux Software Components

Components	Description
OpenFabrics core and ULPs	• IB HCA drivers (mlx4, mlx5) • core • Upper Layer Protocols: IPoIB, SRP and iSER Initiator
OpenFabrics utilities	• OpenSM: IB Subnet Manager with Mellanox proprietary Adaptive Routing • Diagnostic tools • Performance tests
MPI	• OSU MPI (mvapich2-1.9-1) stack supporting the InfiniBand interface • Open MPI stack 1.6.5 and later supporting the InfiniBand interface • MPI benchmark tests (OSU benchmarks, Intel MPI benchmarks, Presta)
PGAS	• ScalableSHMEM v2.2 supporting InfiniBand, MXM and FCA • ScalableUPC v2.2 supporting InfiniBand, MXM and FCA
HPC Acceleration packages	• Mellanox MXM v3.0 (p2p transport library acceleration over Infiniband) • Mellanox FCA v2.5 (MPI/PGAS collective operations acceleration library over InfiniBand) • KNEM, Linux kernel module enabling high-performance intra-node MPI/PGAS communication for large messages
Extra packages	• ibutils2 • ibdump • MFT
Sources of all software modules (under conditions mentioned in the modules' LICENSE files) except for MFT, OpenSM plugins, ibutils2, and ibdump	
Documentation	

4 Supported Platforms and Operating Systems

The following are the supported OSs in MLNX_OFED Rev 2.2-1.0.1:

Table 2 - Supported Platforms and Operating Systems

Operating System	Platform
RHEL/CentOS 6.3	x86_64
RHEL/CentOS 6.4	x86_64/PPC
RHEL/CentOS 6.5	x86_64/PPC
RHEL/CentOS 7.0	x86_64
SLES11 SP1	x86_64
SLES11 SP2	x86_64/PPC
SLES11 SP3	x86_64/PPC
OEL 6.3	x86_64
OEL 6.4	x86_64
OEL 6.5	x86_64
Citrix XenServer Host 6.2	i686
Fedora 19	x86_64
Ubuntu 12.04	x86_64
Ubuntu 13.04	x86_64
Ubuntu 14.04	x86_64
Debian 6.0.7	x86_64
Debian 7.1	x86_64
Debian 7.2	x86_64
kernel 3.10.28[a]	
kernel 3.11.10[a]	
kernel 3.12.9[a]	
kernel 3.13.1[a]	

a. This kernel is supported only when using the Operating Systems stated in the table above.

If you wish to install OFED on a different kernel, you need to create a new ISO image, using mlnx_add_kernel_support.sh script.
See the MLNX_OFED User Guide for instructions.

Upgrading MLNX_OFED on your cluster requires upgrading all of its nodes to the newest version as well.

4.1 Supported Hypervisors

The following are the supported hypervisors in MLNX_OFED Rev 2.2-1.0.1:

- KVM
- Xen

5 Hardware and Software Requirements

The following are the hardware and software requirements of MLNX_OFED Rev 2.2-1.0.1.

- Linux operating system
- Administrator privileges on your machine(s)
- Disk Space: 1GB

For the OFED Distribution to compile on your machine, some software packages of your operating system (OS) distribution are required.

To install the additional packages, run the following commands per OS:

Table 3 - Additional Software Packages

Operating System	Required Packages Installation Command
RHEL/OEL/ Fedora	yum install pciutils python gcc-gfortran libxml2-python tcsh libnl.i686 libnl expat glib2 tcl libstdc++ bc tk
XenServer	yum install pciutils python libxml2-python libnl expat glib2 tcl bc libstdc++ tk
SLES 11 SP1	zypper install pciutils python libxml2-python tcsh libnl libstdc++43 expat glib2 tcl bc libnl.i586 tk
SLES 11 SP2	zypper install pciutils python libnl-32bit libxml2-python tcsh libnl libstdc++46 expat glib2 tcl bc tk
SLES 11 SP3	zypper install pciutils python libnl-32bit libxml2-python tcsh libstdc++43 libnl expat glib2 tcl bc tk
Ubuntu/Debian	apt-get install dpkg autotools-dev autoconf libtool automake1.10 automake m4 dkms debhelper tcl tcl8.4 chrpath swig graphviz tcl-dev tcl8.4-dev tk-dev tk8.4-dev bison flex dpatch zlib1g-dev curl libcurl4-gnutls-dev python-libxml2 libvirt-bin libvirt0 libnl-dev libglib2.0-dev libgfortran3

6 Supported HCAs

MLNX_OFED Rev 2.2-1.0.1 supports the following Mellanox network adapter cards:

- Connect-IB® (Rev 10.10.3000 and above)
- ConnectX®-3 Pro (Rev 2.31.5050 and above)
- ConnectX®-3 (Rev 2.31.5050 and above)
- ConnectX®-2 (Rev 2.9.1200 and above)[1]

For official firmware versions please see:
http://www.mellanox.com/content/pages.php?pg=firmware_download

1. ConnectX®-2 does not support all the new functionality of MLNX_OFED 2.0.3-XXX. For the complete list of the supported features per HCA, please refer to the MLNX_OFED User Manual.

7 Compatibility

MLNX_OFED Rev 2.2-1.0.1 is compatible with the following:

Table 4 - MLNX_OFED Rev 2.2-1.0.1 Compatibility Matrix

Mellanox Product	Description/Version
MLNX-OS®	MSX6036 w/w MLNX-OS® version 3.3.4304[a]
Grid Director™	4036 w/w Grid Director™ version 3.9.1-985
FabricIT™ EFM	IS5035 w/w FabricIT EFM version 1.1.3000
FabricIT™ BXM	MBX5020 w/w FabricIT BXM version 2.1.2000
Unified Fabric Manager (UFM®)	v4.6
MXM	v3.0
ScalableUPC	v2.2
ScalableSHMEM	v2.2
FCA	v2.5
OMPI	v1.6.5
MVAPICH	v1.9a
CD	v1.0

a. MLNX_OFED v2.2-1.0.1 was tested with this switch however, additional switches might be supported as well.

8 Change Log History

8.1 Changes in Rev 2.2-1.0.1 From Rev 2.1-1.0.6

Table 5 - New Features, Changes and Fixes in v2.2-1.0.1

Category	Description
mlnxofedinstall	32-bit libraries are no longer installed by default on 64-bit OS. To install 32-bit libraries use the `'--with-32bit'` installation parameter.
openibd	Added pre/post start/stop scripts support. For further information, please refer to section *"openibd Script"* in the MLNX_OFED User Manual.
Reset Flow	Reset Flow is not activated by default. It is controlled by the mlx4_core`'internal_err_reset'` module parameter.
InfiniBand Core	Asymmetric MSI-X vectors allocation for the SR-IOV hypervisor and guest instead of allocating 4 default MSI-X vectors. The maximum number of MSI-X vectors is `num_cpu` for port ConnectX®-3 has 1024 MSI-X vectors, 28 MSI-X vectors are reserved. • Physical Function - gets the number of MSI-X vectors according to the `pf_msix_table_size` (multiple of 4 - 1) INI parameter • Virtual Functions – the remaining MSI-X vectors are spread equally between all VFs, according to the `num_vfs mlx4_core` module parameter
Ethernet	Ethernet VXLAN support for kernels 3.12.10 or higher
	Power Management Quality of Service: when the traffic is active, the Power Management QoS is enabled by disabling the CPU states for maximum performance.
	Ethernet PTP Hardware Clock support on kernels/OSes that support it
Verbs	Added additional experimental verbs interface. This interface exposes new features which are not integrated yet in to the upstream libibverbs. The Experimental API is an extended API therefor, it is backward compatible, meaning old application are not required to be recompiled to use MLNX-OFED v2.2-1.0.1.
Performance	Out of the box performance improvements: • Use of affinity hints (based on NUMA node of the device) to indicate the IRQ balancer daemon on the optimal IRQ affinity • Improvement in buffers allocation schema (based on the hint above) • Improvement in the adaptive interrupt moderation algorithm

8.2 Changes in Rev 2.1-1.0.6 From Rev 2.1-1.0.0

Table 6 - New Features, Changes and Fixes in v2.1-1.0.6

Category	Description
IB Core	Added allocation success verification process to ib_alloc_device.
dapl	dapl is recompiled with no FCA support.

Table 6 - New Features, Changes and Fixes in v2.1-1.0.6

Category	Description
openibd	Added the ability to bring up child interfaces even if the parent's ifcfg file is not configured.
libmlx4	Unmapped the hca_clock_page parameter from mlx4_uninit_context.
scsi_transport_srp	scsi_transport_srp cannot be cleared up when rport reconnecting fails.
mlnxofedinstall	Added support for the following parameters: • '--umad-dev-na' • '--without-<package>'
Content Packages Updates	The following packages were updated: • bupc to v2.2-407 • mstflint to v3.5.0-1.1.g76e4acf • perftest to v2.0-0.76.gbf9a463 • hcoll to v2.0.472-1 • Openmpi to v1.6.5-440ad47 • dapl to v2.0.40

8.3 Changes in Rev 2.1-1.0.0 From Rev 2.0-3.0.0

Table 7 - New Features, Changes and Fixes in v2.1-1.0.0

Category	Description
EoIB	EoIB is supported only in SLES11SP2 and RHEL6.4.
eIPoIB	eIPoIB is currently at GA level.
Connect-IB®	Added the ability to resize CQs.
IPoIB	Reusing DMA mapped SKB buffers: Performance improvements when IOMMU is enabled.
mlnx_en	Added reporting autonegotiation support.
	Added Transmit Packet Steering (XPS) support.
	Added reporting 56Gbit/s link speed support.
	Added Receive Flow Steering (RFS) support in UDP.
	Added Low Latency Socket (LLS) support.
	Added check for dma_mapping errors.
eIPoIB	Added non-virtual environment support.
Hypervisor support	KVM and XenServer.

8.4 Changes in Rev 2.0-3.0.0 From Rev 2.0-2.0.5

Table 8 - New Features, Changes and Fixes in v2.0-3.0.0

Category	Description
Operating Systems	Additional OS support: • SLES11SP3 • Fedora16, Fedora17
Drivers	Added Connect-IB™ support
Installation	Added ability to install MLNX_OFED with SR-IOV support.
	Added Yum installation support
EoIB	EoIB (at beta level) is supported only in SLES11SP2 and RHEL6.4
mlx4_core	Modified module parameters to associate configuration values with specific PCI devices identified by their bus/device/function value format
mlx4_en	Reusing DMA mapped buffers: major performance improvements when IOMMU is enabled
	Added Port level QoS support
IPoIB	Reduced memory consumption
	Limited the number TX and RX queues to 16
	Default IPoIB mode is set to work in Datagram, except for Connect-IB™ adapter card which uses IPoIB with Connected mode as default.
Storage	iSER (at GA level)

8.5 New Features in Rev 2.0-2.0.5[1]

- SR-IOV for both Ethernet and InfiniBand (at Beta level)
- RoCE over SR-IOV (at Beta level)
- eIPoIB to enable IPoIB in a Para-Virtualized environment (at Alpha level)
- Contiguous pages:
  - Internal memory allocation improvements
  - Register shared memory
  - Control objects (QPs, CQs)
- Ethernet Performance Enhancements (NUMA related and others) for 10G and 40G
- OFED_VMA integration to a single branch
- Ethernet Time Stamping (at Beta level)
- Flow Steering for Ethernet and InfiniBand. (at Beta level)
- Raw Eth QPs:
  - Checksum TX/RX

1. SR-IOV, Ethernet Time Stamping and Flow Steering are ConnectX®-3 HCA capability.

  - Flow Steering
- Errata Kernel upgrade support
- YUM update support
- Storage – iSER (at Beta level) and SRP
- 64bit wide counters (port xmit/recv data/packets unicast/mcast)
- VERSION query API: library and headers

9 Known Issues

The following is a list of general limitations and known issues of the various components of this Mellanox OFED for Linux release.

Table 9 - Known Issues

Index	Issue	Description	Workaround
1.	IPoIB	When user increases receive/send a buffer, it might consume all the memory when few child's interfaces are created.	-
2.		The hardware address suffix of IPoIB interfaces in MLNX_OFED v2.0-3.0.0 is `'a'` instead of `'8'` to indicate the TSS support.	-
3.		The size of send queue in Connect-IB® cards cannot exceed 1K.	-
4.		In 32 bit devices, the maximum number of child interfaces that can be created is 16. Creating more that, might cause out-of-memory issues.	-
5.		The default IPoIB operating mode in ConnectX® family adapter cards is UD and CM in Connect-IB®.	-
6.		Changing the IPoIB mode (CM vs UD) requires the interface to be in `'down'` state.	-
7.		In RHEL7.0, the Network-Manager can detect when the carrier of one of the IPoIB interfaces is OFF and can decide to disable its IP address.	-
8.		IPoIB interface does not function properly if a third party application changes the PKey table. We recommend modifying PKey tables via OpenSM.	-
9.		When creating a new child interface in an over-loaded kernel, a `dmesg` print is displayed advising the user to try again in a few seconds.	-
10.		Out-of memory issue might occur due to overload of interfaces created.	To calculate the allowed memory per each IPoIB interface check the following: • Num-rings = min(num-cores-on-that-device, 16) • Ring-size = 512 (by default, it is module parameter) • UD memory: 2 * num-rings * ring-size * 8K • CM memory: ring-size * 64k • Total memory = UD mem + CM mem

Table 9 - Known Issues (Continued)

Index	Issue	Description	Workaround
11.		The physical port MTU (indicates the port capability) default value was changed to 4k, whereas the IPoIB port MTU ("logical" MTU) default value is 2k as it is set by the OpenSM.	In order to change the IPoIB MTU to 4k, edit the OpenSM partition file in the section of IPoIB setting as follow: `Default=0xffff, ipoib, mtu=5 : ALL=full;` *Where `"mtu=5"` indicates that all IPoIB ports in the fabric are using 4k MTU, (`"mtu=4"` indicates 2k MTU)
12.		Occasionally, when using IPoIB in Connected mode, the connection might get closed and recovered only after several minutes.	Use the Datagram mode
13.		Fallback to the primary slave of an IPoIB bond does not work with ARP monitoring. (https://bugs.openfabrics.org/show_bug.cgi?id=1990)	-

14.		Whenever the IOMMU parameter is enabled in the kernel it can decrease the number of child interfaces on the device according to resource limitation.
		The driver will stuck after unknown amount of child interfaces creation.
		
		For further information, please see:
		https://access.redhat.com/site/articles/66747
		http://support.citrix.com/article/CTX136517
		http://www.novell.com/support/kb/doc.php?id=7012337
			To avoid such issue:
			• Decrease the amount of the RX receive buffers (module parameter, the default is 512)
			• Decrease the number of RX rings (sys/fs or ethtool in new kernels)
			• Avoid using IOMMU if not required
			
			For **KVM** users:
			Run:
			`echo 1 > /sys/module/kvm/parameters/allow_unsafe_assigned_interrupts`
			
			To make this change persist across reboots, add the following to the `/etc/modprobe.d/kvm.conf` file (or create this file, if it does not exist):
			`options kvm allow_unsafe_assigned_interrupts=1 kernel parameters`

Table 9 - Known Issues (Continued)

Index	Issue	Description	Workaround

15.		System might crash in `skb_checksum_help()` while performing TCP retransmit involving packets with 64k packet size.
		A similar out to the below will be printed:
		`kernel BUG at net/core/dev.c:1707!`
		`invalid opcode: 0000 [#1] SMP`
		`RIP: 0010:[<ffffffff81448988>]`
		`skb_checksum_help+0x148/0x160`
		`Call Trace:`
		`<IRQ>`
		`[<ffffffff81448d83>]`
		`dev_hard_start_xmit+0x3e3/0x530`
		`[<ffffffff8144c805>]`
		`dev_queue_xmit+0x205/0x550`
		`[<ffffffff8145247d>]`
		`neigh_connected_output+0xbd/0x1`
		…
		…
			Use UD mode in ipoib

16.		Changing the GUID of a specific SR-IOV guest after the driver has been started, causes the ping to fail. Hence, no traffic can go over that InfiniBand interface.	-
17.		`send_queue_size` over Connect-IB™ adapter cards cannot be larger than 1024	-
18.	Ethernet	Ethernet PV VLAN Guest transparent Tagging (VGT) is only supported in openvswitch and not in standard Linux vBridges and libvirt For more information please see : http://libvirt.org/formatnetwork.html (Setting VLAN tag section)	-

19.		Changing the ring size on 32-bit system may result in failure due to lack of memory. Therefore, mlx4_en will not be able to vmap enough memory and the below message will be printed in dmesg:
		`vmap allocation for size 528384 failed: use vmalloc=<size> to increase size`
		`In this case user can enlarge the vmalloc memory by adding vmalloc=<size> to grub.conf`
		`Default vmalloc setting is 128M. It is recommended to add each time 64M of memory until desired ring size can be allocated.`
		Please note, that in case vmalloc size is too big, the OS will fail to boot, so please use caution when adding additional memory.
		For more info refer to: http://www.mythtv.org/wiki/Common_Problem:_vmalloc_too_small
			-

Table 9 - Known Issues (Continued)

Index	Issue	Description	Workaround
20.		On OEL6.1 with uek1 (2.6.32-x.x.x.el6uek kernel), when the number of RX ring is smaller than TX rings (kernel issue), the following call trace will be shown in the kernel log: `WARNING: at net/core/dev.c:2077 get_rps_cpu+0x70/0x2b9()`	-
21.		Kernel panic might occur during traffic over IPv6 on kernels between 3.12-rc7 and 3.13-rc1 (kernel issue)	-
22.		Kernel panic might occur during fio splice in kernels before 2.6.34-rc4.	Use kernel v2.6.34-rc4 which provides the following solution: `baff42a net: Fix oops from tcp_collapse() when using splice()`
23.		When creating more than 125 VLANs and SR-IOV mode is enabled, a kernel warning message will be printed indicating that the native VLAN is created but will not work with RoCE traffic. kernel warning: `mlx4_core 0000:07:00.0: vhcr command ALLOC_RES (0xf00) slave:0 in_param 0x7e in_mod=0x107, op_mod=0x1 failed with error:0, status -28`	-
24.		In PPC systems when QoS is enabled a harmless Kernel DMA mapping error messages might appear in kernel log (iommu related issue).	-
25.		On kernels that do not support multi-queues, the number of TX channels represents the number of TX rings. The maximal number of TX channels is 16.	-
26.		Transmit timeout might occur on RH6.3 as a result of lost interrupt (OS issue). In this case, the following message will be shown in dmesg: `do_IRQ: 0.203 No irq handler for vector (irq -1)`	-

Table 9 - Known Issues (Continued)

Index	Issue	Description	Workaround
27.	eIPoIB	On rare occasions, upon driver restart the following message is shown in the dmesg: `'cannot create duplicate filename '/class/net/eth_ipoib_interfaces'`	-
28.		No indication is received when eIPoIB is non functional.	Run `'ps -ef | grep ipoibd'` to verify its functionality.
29.		eIPoIB requires libvirtd, python	-
30.		eIPoIB supports only active-backup mode for bonding.	-
31.		eIPoIB supports only VLAN Switch Tagging (VST) mode on guests.	-
32.		IPv6 is currently not supported in eIPoIB	-
33.		eIPoIB cannot run when Flow Steering is enabled	-

34.	XRC	Legacy API is deprecated, thus when recompiling applications over MLNX_OFED v2.0-3.x.x, warnings such as the below are displayed.
		`rdma.c:1699: warning: 'ibv_open_xrc_domain' is deprecated (declared at /usr/include/infiniband/ofa_verbs.h:72)`
		`rdma.c:1706: warning: 'ibv_create_xrc_srq' is deprecated (declared at /usr/include/infiniband/ofa_verbs.h:89)`
		These warnings can be safely ignored.
			-

35.		XRC is not functional in heterogeneous clusters containing non Mellanox HCAs.	-
36.		XRC options do not work when using qperf tool.	Use perftest instead
37.		Out-of memory issue might occur due to overload of XRC receive QP with non zero receive queue size created. XRC QPs do not have receive queues.	-
38.	mlx4_ib module	The `dev_assign_str` module parameter is not backward compatible. In the current version, this parameter is using decimal number to describe the InfiniBand device and not hexadecimal number as it was in previous versions in order to uniform the mapping of device function numbers to InfiniBand device numbers as defined for other module parameters (e.g. `num_vfs` and `probe_vf`).	-
39.	ABI Compatibility	MLNX_OFED Rev 2.2-1.0.1 is not ABI compatible with previous MLNX_OFED/OFED versions.	Recompile the application over the new MLNX_OFED version

Table 9 - Known Issues (Continued)

Index	Issue	Description	Workaround
40.	System Time	Loading the driver using the openibd script when no InfiniBand vendor module is selected (for example mlx4_ib), may cause the execution of the /sbin/start_udev' script. In RedHat 6.x and OEL6.x this may change the local system time.	-
41.	ConnectX-3 Adapter Cards Family	Using RDMA READ with a higher value than 30 SGEs in the WR might lead to "local length error".	Do not set the value of SGEs higher than 30 when RDMA READ is used.
42.	Verbs	Using libnl1_1_3~26 or earlier, requires ibv_create_ah protection by a lock for multi-threaded applications.	-
43.	Driver Start	"Out of memory" issues may rise during drivers load depending on the values of the driver module parameters set (e.g. log_num_cq).	-
44.		When reloading/starting the driver using the `/etc/init.d/openibd` the following messages are displayed if there is a third party RPM or driver installed: "Module mlx4_core does not belong to MLNX_OFED" or "Module mlx4_core belong to <rpm name> which is not a part of MLNX_OFED"	Remove the third party RPM/non MLNX_OFED drivers directory, run: `"depmod"` and then rerun `"/etc/init.d/openibd restart"`
45.		Occasionally, when trying to repetitively reload the nes hardware driver on SLES11 SP2, a soft lockups occurs that required reboot.	-
46.		In ConnectX-2, if the driver load succeeds, the informative message below is presented conveying the below limitations: • If port type is IB the number of maximum supported VLs is 4 • If port type is ETH then the maximum priority for VLAN tagged is 3 `"mlx4_core 0000:0d:00.0: command SET_PORT (0xc) failed: in_param=0x120064000, in_mod=0x2, op_mod=0x0, fw status = 0x40"`	-
47.	Performance Tools	perftest package in MLNX_OFED v2.2-1.0.1 does not work with older versions of the driver.	-

Table 9 - Known Issues (Continued)

Index	Issue	Description	Workaround
48.	SR-IOV	When using legacy VMs with OFED 2.0-2.0.5 hypervisor, the `'enable_64b_cqe_eqe'` parameter must be set to zero on the hypervisor. It should be set in the same way that other module parameters are set for mlx4_core at module load time. For example, add `"options mlx4_core enable_64b_cqe_eqe=0"` as a line in the file `/etc/modprobe.d/mlx4_core.conf`.	-
49.		When working with a bonding device to enslave the Ethernet devices in active-backup mode and failover MAC policy in a Virtual Machine (VM), establishment of RoCE connections may fail.	Unload the module mlx4_ib and reload it in the VM.
50.		Trying to establish RoCE connection with RDMA_CM may fail in a VM.	Unload the module mlx4_ib and reload it in the VM.
51.		When the number of single port VFs is not the same for port 1 and port 2 and the relevant VF is single ported on port 2, RoCE connections with RDMA_CM will fail in a VM.	-
52.		When detaching a VF without shutting down the driver from a VM and reattaching it to another VM with the same IP address for the Mellanox NIC, RoCE connections will fail	Shut down the driver in the VM before detaching the VF.
53.		Enabling SR-IOV requires appending the `"intel_iommu=on"` option to the relevant OS in file `/boot/grub/grub.conf/`. Without that SR-IOV cannot be loaded.	-
54.		rdma_cm does not support UD QPs	-
55.		SR-IOV can be enabled only when using the firmware version embedded in the MLNX_OFED v2.0-3.0.0 and above.	-
56.		On various combinations of Hypervisor/OSes and Guest/OSes, an issue might occur when attaching/ detaching VFs to a guest while that guest is up and running.	Attach/detach VFs to/from a VM only while that VM is down.
57.		When SR-IOV is disabled in the system BIOS, a PCI issue is noticed in Ubuntu v12.04.3 with Linux kernel v3.8 which affects NICs of several manufacturers including Mellanox's, preventing them from operating.	Enable SR-IOV in the BIOS
58.		When working with SR-IOV in Xen-4.2 virtualization platform, only the built-in xen_pciback driver should be loaded. The xen_pciback module in dom0 should not be loaded, as loading them simultaneously may cause interrupts loss and cause the driver to enter the reset flow.	-

Table 9 - Known Issues (Continued)

Index	Issue	Description	Workaround
59.		The known PCI BDFs for all VFs in kernel command line should be specified by adding `xen-pciback.hide` For further information, please refer to http://wiki.xen.org/wiki/Xen_PCI_Passthrough	-
60.		mlx4_core module parameter `"port_type_array"` cannot specify 3 (= auto) as a port type value because of a mistake in the `"modinfo mlx4_core"` description of this parameter.	-
61.	Port Type Management	OpenSM must be stopped prior to changing the port protocol from InfiniBand to Ethernet.	-
62.		After changing port type using `connectx_port_config` interface ports' names can be changed. For example. ib1 -> ib0 if port1 changed to be Ethernet port and port2 left IB.	Use `udev` rules for persistent naming configuration. For further information, please refer to the User Manual
63.	Flow Steering	Flow Steering is disabled by default.	To enable it, set the parameter below as follow: `log_num_mgm_entry_size should set to -1`
64.		IPv4 rule with source IP cannot be created in SLES 11	-
65.		RFS is not supported in SLES11	-
66.	Quality of Service	QoS is not supported in XenServer, Debian 6.0 and 6.2 with uek kernel	-
67.	Driver Uninstall	A Kernel panic occurs if you uninstall the driver without deleting the SR-IOV module params (`mlx4_core's num_vfs`) in the file `/etc/modprobe.d/mlx4_core.conf`. On the next boot, you will get the panic, and machine will boot up.	Remove the midule after uninstalling and prior to restarting the driver.
68.	Installation	When upgrading from an earlier Mellanox OFED version, the installation script does not stop the earlier version prior to uninstalling it.	Stop the old OFED stack (`/etc/init.d/openibd stop`) before upgrading to this new version.
69.		Upgrading from the previous OFED installation to this release, does not unload the kernel module ipoib_helper.	Reboot after installing the driver.
70.		Installation using Yum does not update HCA firmware.	See "Updating Firmware After Installation" in OFED User Manual
71.		On SLES11.1 the package 'libnl.i586' is required to install MLNX_OFED.	Perform one of the following: • Install the 'libnl.i586' RPM from the SLES11.1 32bit installation disk • Install MLNX_OFED with the following flag "--without-32bit"

Table 9 - Known Issues (Continued)

Index	Issue	Description	Workaround
72.		"--total-vfs <0-63>" installation parameter is no longer supported	Use '--enable-sriov' installation parameter to burn firmware with SR-IOV support. The number of virtual functions (VFs) will be set to 16. For further information, please refer to the User Manual.
73.		When using bonding on Ubuntu OS, the "ifenslave" package must be installed.	-
74.	Driver Unload	"openibd stop" can sometime fail with the error: Unloading ib_cm [FAILED] ERROR: Module ib_cm is in use by ib_ipoib	Re-run "openibd stop"
75.	Fork Support	Fork support from kernel 2.6.12 and above is available provided that applications do not use threads. fork() is supported as long as the parent process does not run before the child exits or calls exec(). The former can be achieved by calling wait(childpid), and the latter can be achieved by application specific means. The Posix system() call is supported.	-
76.	ISCSI over IPoIB	When working with ISCSI over IPoIB, LRO must be disabled (even if IPoIB is set to connected mode) due to a a bug in older kernels which causes a kernel panic.	-
77.	MLNX_OFED sources	MLNX_OFED includes the OFED source RPM packages used as a build platform for kernel code but does not include the sources of Mellanox proprietary packages.	-
78.	InfiniBand Utilities	When running the ibdiagnet check nodes_info on the fabric, a warning specifying that the card does not support general info capabilities for all the HCAs in the fabric will be displayed.	Run ibdiagnet --skip nodes_info
79.	mlx5 Driver	Atomic Operations in Connect-IB® are fully supported on big-endian machines (e.g. PPC). Their support is limited on little-endian machines (e.g. x86)	-
80.	General	On ConnectX-2/ConnectX-3 Ethernet adapter cards, there is a mismatch between the GUID value returned by firmware management tools and that returned by fabric/driver utilities that read the GUID via device firmware (e.g., using ibstat). Mlxburn/flint return 0xffff as GUID while the utilities return a value derived from the MAC address. For all driver/firmware/software purposes, the latter value should be used.	N/A. Please use the GUID value returned by the fabric/driver utilities (not 0xffffff).

Table 9 - Known Issues (Continued)

Index	Issue	Description	Workaround
81.	Uplinks	On rare occasions, ConnectX®-3 Pro adapter card may fail to link up when performing parallel detect to 40GbE.	Restart the driver
82.	Resources Limitation	The device capabilities reported may not be reached as it depends on the system on which the device is installed and whether the resource is allocated in the kernel or the userspace.	-

83.		Occasionally, a user process might experience some memory shortage and not function properly due to Linux kernel occupation of the system's free memory for its internal cache.	To free memory to allow it to be allocated in a user process, run the `drop_caches` procedure below.
			Performing the following steps will cause the kernel to flush and free pages, dentries and inodes caches from memory, causing that memory to become free.
			
			Note: As this is a non-destructive operation and dirty objects are not freeable, run `` `sync' `` first.
			
			• To free the pagecache:
			`echo 1 > /proc/sys/vm/drop_caches`
			• To free dentries and inodes:
			`echo 2 > /proc/sys/vm/drop_caches`
			• To free pagecache, dentries and inodes:
			`echo 3 > /proc/sys/vm/drop_caches`

84.		Setting more IP addresses than the available GID entries in the table results in failure and the "update_gid_table error message is displayed: `GID table of port 1 is full. Can't add <address>" message.`	-
85.	Ethernet Performance Counters	In a system with more than 61 VFs, the 62nd VF and onwards is assigned with the SINKQP counter, and as a result will have no statistics, and loopback prevention functionality for SINK counter.	-
86.		Since each VF tries to allocate 2 more QP counter for its RoCE traffic statistics, in a system with less than 61 VFs, if there is free resources it receives new counter otherwise receives the default counter which is shared with Ethernet. In this case RoCE statistics is not available.	-

Table 9 - Known Issues (Continued)

Index	Issue	Description	Workaround
87.		In ConnectX®-3, when we enable function-based loopback prevention for Ethernet port by default (i.e., based on the QP counter index), the dropped self-loopback packets increase the IfRxError-Frames/Octets counters.	-
88.	RoCE	Not configuring the Ethernet devices or independent VMs with a unique IP address in the physical port, may result in RoCE GID table corruption.	Restart the driver
89.		If RDMA_CM is not used for connection management, then the source and destination GIDs used to modify a QP or create AH should be of the same type - IPv4 or IPv6.	-
90.		On rare occasions, the driver reports a wrong GID table (read from /sys/class/infiniband/mlx4_*/ports/*/gids/*). This may cause communication problems.	-
91.		Since the number of GIDs per port is limited to 128, there cannot be more than the allowed IP addresses configured to Ethernet devices that are associated with the port. Allowed number is: • "127" for a single function machine • "15" for a hypervisor in a multifunction machine • "n" for a guest in a multifunction machine (where n is the number of virtual functions)	-
92.		A working IP connectivity between the RoCE devices is required when creating an address handle or modifying a QP with an address vector.	-
93.		MLNX_OFED v2.1-1.0.0 is not interoperable with older versions of MLNX_OFED.	-
94.		When configuring the single-port-per-VF feature, if the number of VFs on port 1 is different from port 2, RoCE/rdma_cm on single-ported VFs which use port 2 may not work.	-
95.		Unloading mlx4_en while a rdma_cm session is established can cause a kernel panic.	-
96.	Storage	SLES11-SP1: When running multipath rescan while new devices are added to mpath tables, multipath may not find all the device-mappers.	-
97.		Older versions of rescan_scsi_bus.sh may not recognize some newly created LUNs.	If encountering such issues, it is recommended to use the '-c' flag.
98.	SRP	Reconnecting to a target during host reset stage may result in devices going Offline.	Run `rescan-scsi-bus.sh -r`
99.		MLNX_OFED SRP installation breaks the ibmvstgt and ibmvscsi symbol resolution in RHEL7.0	-
100.	SRP Interop	The driver is tested with Storage target vendors recommendations for multipath.conf extensions (ZFS, DDN, TMS, Nimbus, NetApp).	-

Table 9 - Known Issues (Continued)

Index	Issue	Description	Workaround
101.	DDN Storage Fusion 10000 target	DDN does not accept non-default P_Key connection establishment.	-
102.	Oracle Sun ZFS storage 7420	Occasionally the first command to a LUN may not be serviced, aborted, and cause a successful re-connection to the target	-
103.		ZFS does not accept non-default P_Key connection establishment.	-
104.		Ungraceful power cycle of an initiator connected with Targets DDN, Nimbus, NetApp may result in temporary "stale connection" messages when initiator reconnects.	-
105.	iSER	On SLES11, the `ib_iser` module does not load on boot	Add a dummy interface using iscsiadm: • `# iscsiadm -m iface -I ib_iser -o new` • `# iscsiadm -m iface -I ib_iser -o update -n iface.transport_name -v ib_iser`
106		In Ubuntu12.04 need to update user space open-iscsi package to version 2.0.873.	-
107		Trying to disconnect a session while the session is undergoing a reconnect flow may result in disconnection hang.	Restart iscsid. **Note:** Please be aware that doing so might cause `rmmod` process to hang as the ib_iser module will not be unloaded.
108		Unloading ib_iser during session disconnect event may result in kernel panic.	-
109		Ubuntu12 requires update of user space open-iscsi to v2.0.873	-
110		The initiator does not respect interface parameter while logging in.	Configure each interface on a different subnet.
111		When connecting to a target with no exposed LUNs, the following message may be reported `"Unexpected RDMA CM event"`. The message can be safely ignored.	-
112		The initiator may take a long time to stop trying to log in via a non RDMA capable interface.	-
113	ZFS appliance	Connection establishment occurs twice which may cause iSER to log a stack trace.	-
114	iSER interop - Oracle Sun ZFS storage 7420	Connection establishment occurs twice which may cause iSER to log a stack trace	-

10 API Changes

10.1 API Changes in MLNX_OFED Rev 2.2-1.0.1

The following are the API changes in MLNX_OFED Rev 2.2-1.0.1:

Table 10 - API Changes in MLNX_OFED Rev 2.2-1.0.1

Name	Description
libibverbs	The following verbs changed to align with upstream libibverbs: • `ibv_reg_mr - ibv_access_flags changed.` • `ibv_post_send - opcodes and send flags changed and wr fields removed (task, op, dc and bind_mw)` • `ibv_query_device - capability flags changed.` • `ibv_poll_cq - opcodes and wc flags changed.` • `ibv_modify_qp - mask bits changed` • `ibv_create_qp_ex - create_flags field removed.` The following verbs removed to align with upstream libibverbs: • `ibv_bind_mw` • `ibv_post_task` • `ibv_query_values_ex` • `ibv_query_device_ex` • `ibv_poll_cq_ex` • `ibv_reg_shared_mr_ex` • `ibv_reg_shared_mr` • `ibv_modify_cq` • `ibv_create_cq_ex` • `ibv_modify_qp_ex`
Verbs Experimental API	The following experimental verbs added (replacing the removed extended verbs): • ibv_exp_bind_mw • ibv_exp_post_task • ibv_exp_query_values • ibv_exp_query_device • ibv_exp_poll_cq • ibv_exp_reg_shared_mr • ibv_exp_modify_cq • ibv_exp_create_cq • ibv_exp_modify_qp New experimental verbs: • ibv_exp_arm_dct • ibv_exp_query_port • ibv_exp_create_flow • ibv_exp_destroy_flow • ibv_exp_post_send • ibv_exp_reg_mr • ibv_exp_get_provider_func

10.2 API Changes in MLNX_OFED Rev 2.1-1.0.0

The following are the API changes in MLNX_OFED Rev 2.1-1.0.0:

Table 11 - API Changes in MLNX_OFED Rev 2.1-1.0.0

Name	Description
Dynamically Connected (DC)	The following verbs were added: • `struct ibv_dct *ibv_exp_create_dct(struct ibv_context *context, struct ibv_exp_dct_init_attr *attr)` • `int ibv_exp_destroy_dct(struct ibv_dct *dct)` • `int ibv_exp_query_dct(struct ibv_dct *dct, struct ibv_exp_dct_attr *attr)`
Verbs Extension API: Verbs extension API defines OFA APIs extension scheme to detect ABI compatibility and enable backward and forward compatibility support.	• ibv_post_task • ibv_query_values_ex • ibv_query_device_ex • ibv_create_flow • ibv_destroy_flow • ibv_poll_cq_ex • ibv_reg_shared_mr_ex • ibv_open_xrcd • ibv_close_xrcd • ibv_modify_cq • ibv_create_srq_ex • ibv_get_srq_num • ibv_create_qp_ex • ibv_create_cq_ex • ibv_open_qp • ibv_modify_qp_ex
Verbs Experimental API: Verbs experimental API defines MLNX-OFED APIs extension scheme which is similar to the “Verbs extension API”. This extension provides a way to introduce new features before they are integrated into the formal OFA API and to the upstream kernel and libs.	• ibv_exp_create_qp • ibv_exp_query_device • ibv_exp_create_dct • ibv_exp_destroy_dct • ibv_exp_query_dct

10.3 API Changes in MLNX_OFED Rev 2.0-3.0.0

The following are the API changes in MLNX_OFED Rev 2.0-3.0.0:

Table 12 - API Changes in MLNX_OFED Rev 2.0-3.0.0

Name	Description
XRC	The following verbs have become deprecated: • `struct ibv_xrc_domain *ibv_open_xrc_domain` • `struct ibv_srq *ibv_create_xrc_srq` • `int ibv_close_xrc_domain` • `int ibv_create_xrc_rcv_qp` • `int ibv_modify_xrc_rcv_qp` • `int ibv_query_xrc_rcv_qp` • `int ibv_reg_xrc_rcv_qp` • `int ibv_unreg_xrc_rcv_qp`

10.4 API Changes in MLNX_OFED Rev 2.0-2.0.5

The following are the API changes in MLNX_OFED v2.0-2.0.5:

Table 13 - API Changes in MLNX_OFED Rev 2.0-2.0.5

Name	Description
Libibverbs	
Extended speeds	• Missing the `ext_active_speed attribute` from the struct `ibv_port_attr` • Removed function `ibv_ext_rate_to_int` • Added functions `ibv_rate_to_mbps` and `mbps_to_ibv_rate`
Raw QPs	QP types `IBV_QPT_RAW_PACKET` and `IBV_QPT_RAW_ETH` are not supported
Contiguous pages	• Added Contiguous pages support • Added function `ibv_reg_shared_mr`
Libmverbs	
	• The enumeration `IBV_M_WR_CALC` was renamed to `IBV_M_WR_CALC_SEND` • The enumeration `IBV_M_WR_WRITE_WITH_IMM` was added • In the structure `ibv_m_send_wr`, the union `wr.send` was renamed to `wr.calc_send` and `wr.rdma` was added • The enumerations `IBV_M_WQE_CAP_CALC_RDMA_WRITE_WITH_IMM` was added

	• The following enumerations were renamed:
	  • From `IBV_M_WQE_SQ_ENABLE_CAP` to `IBV_M_WQE_CAP_SQ_ENABLE`
	  • From `IBV_M_WQE_RQ_ENABLE_CAP` to `IBV_M_WQE_CAP_RQ_ENABLE`
	  • From `IBV_M_WQE_CQE_WAIT_CAP` to `IBV_M_WQE_CAP_CQE_WAIT`
	  • From `IBV_M_WQE_CALC_CAP` to `IBV_M_WQE_CAP_CALC_SEND`

11 Bug Fixes History

Table 14 lists the bugs fixed in this release.

Table 14 - Fixed Bugs List

#	Issue	Description	Discovered in Release	Fixed in Release
1.	Ethernet	Fixed kernel panic on Debian-6.0.7 which occurred when the number of TX channels was set above the default value.	2.1-1.0.0	2.2-1.0.1
2.		Fixed a crash incidence which occurred when enabling Ethernet Time-stamping and running VLAN traffic.	2.0-2.0.5	2.2-1.0.1
3.	XRC	XRC over ROCE in SR-IOV mode is not functional	2.0-3.1.0	2.2-1.0.1
4.	mlx4_en	Fixed wrong calculation of packet true-size reporting in LRO flow.	2.1-1.0.0	2.2-1.0.1
5.	IB Core	Fixed the QP attribute mask upon smac resolving	2.1-1.0.0	2.1-1.0.6
6.	mlx5_ib	Fixed a send WQE overhead issue	2.1-1.0.0	2.1-1.0.6
7.		Fixed a NULL pointer dereference on the debug print	2.1-1.0.0	2.1-1.0.6
8.		Fixed arguments to kzalloc	2.1-1.0.0	2.1-1.0.6
9.	mlx4_core	Fixed the locks around completion handler	2.1-1.0.0	2.1-1.0.6
10.	mlx4_core	Restored port types as they were when recovering from an internal error.	2.0-2.0.5	2.1-1.0.0
11.		Added an N/A port type to support port_type_array module param in an HCA with a single port	2.0-2.0.5	2.1-1.0.0
12.	SR-IOV	Fixed memory leak in SR-IOV flow.	2.0-2.0.5	2.0-3.0.0
13.		Fixed communication channel being stuck	2.0-2.0.5	2.0-3.0.0
14.	mlx4_en	Fixed ALB bonding mode failure when enslaving Mellanox interfaces	2.0-3.0.0	2.1-1.0.0
15.		Fixed leak of mapped memory	2.0-3.0.0	2.1-1.0.0
16.		Fixed TX timeout in Ethernet driver.	2.0-2.0.5	2.0-3.0.0
17.		Fixed ethtool stats report for Virtual Functions.	2.0-2.0.5	2.0-3.0.0
18.		Fixed an issue of VLAN traffic over Virtual Machine in paravirtualized mode.	2.0-2.0.5	2.0-3.0.0
19.		Fixed ethtool operation crash while interface down.	2.0-2.0.5	2.0-3.0.0
20.	IPoIB	Fixed memory leak in Connected mode.	2.0-2.0.5	2.0-3.0.0
21.		Fixed an issue causing IPoIB to avoid pkey value 0 for child interfaces.	2.0-2.0.5	2.0-3.0.0